Plant RNA Interactome Capture: Revealing the Plant RBPome

Marcel Bach-Pages[1], Alfredo Castello[2] and Gail M. Preston[1]*

[1]Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, UK
[2]Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

***Corresponding author:** gail.preston@plants.ox.ac.uk (Gail M. Preston)

Key words: RNA-binding protein, RNA-binding domain, RBPome, UV crosslinking

The application of RNA interactome capture to plants has enabled comprehensive determination of the plant RNA-binding proteome and identified novel families of RNA-binding proteins (RBPs). The technique is providing insight into the evolution of the eukaryotic repertoire of RBPs and will enhance prospects for engineering RBPs to improve crop traits.

RNA interactome capture

The study of RNA-protein complexes started in the 1950's with the visualization of transcripts associated with proteins following transcription by RNA polymerase II. Since then, many experimental studies have aimed to characterize ribonucleoprotein (RNP) complexes from diverse organisms using different techniques (see Singh *et al.* [1] for a review) leading to the biochemical characterization of a number of RNA-binding proteins (RBPs). For instance, Pramanik and Bewley used a combination of UV crosslinking and chromatography on oligo(dT)-cellulose columns to study RNPs in alfalfa embryogenesis [2]. Other studies purified RNP complexes using velocity sedimentation, gel filtration or formaldehyde fixation and isopycnic centrifugation.

In the recent years, the combination of these techniques with high-throughput technologies has boosted the discovery of RBPs. Thus, the use of fluorescent RNA probes coupled to microarrays or immobilized RNA probes coupled to mass spectrometry has enabled identification of hundreds of RBPs. Computational analyses have also been used to identify RBPs based on protein homology and identification of known RNA-binding domains (RBDs). While valuable to catalogue RBPs harbouring classical RNA-binding domains, these studies failed to identify unorthodox RBPs that lack similarity with known RBPs. Moreover, none of these approaches has succeeded in determining the full range of RBPs that are active under different cellular states.

In 2012 Castello *et al*. and Baltz *et al*., described a technique termed RNA interactome capture that enabled comprehensive identification of the human RNA-binding proteome (RBPome) [3–5]. This technique offers the unprecedented opportunity to study the dynamic behaviour of RBPs under different cellular conditions since it enables global identification of active RBPs *in vivo*. However, because RNA interactome capture is based on the pull-down of poly(A) RNAs crosslinked to RBPs, it possesses some limitations that should be considered. For instance, some RBPs might be under-represented due to low efficiency of UV crosslinking, very transient/weak interaction with RNAs, or binding to non-polyadenylated RNAs (e.g. most of the mature organelle RNAs) rather than poly(A) RNAs. The application of this method has rapidly expanded to several human cell lines; and organisms including *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster*, and the parasites *Trypanosoma brucei*, *Leishmania donovani* and *Plasmodium falciparum*. Therefore, RNA interactome capture has made an important contribution to the RNA field by enabling systematic identification of RBPomes in different biological systems.

Recent studies by Marondedze *et al.* (2016), Reichel *et al.* (2016) and Zhang *et al.* (2016) [6–8] have provided the first experimental evidence that RNA interactome

capture can be applied to plants (Figure 1), enabling the determination, for the first time, of the plant RBPome [9].

RNA interactome capture reveals both known and novel proteins in the plant RBPome

Unsurprisingly, all three studies identified a number of known RBPs or proteins harbouring known RBDs, thus affirming the robustness of the technique. They also identified *in silico* predicted RBPs, demonstrating that they indeed bind RNA *in vivo.* A subset of the RBPs identified have orthologs in other organisms including yeast, mouse or humans, and thus are likely to represent conserved or core RBPs. These core proteins are normally associated with translation or splicing.

The three studies identified 79 RBPs as the core RBPome active in the different *Arabidopsis* materials tested, including cell suspension cultures and leaves [6]; etiolated seedlings [7], and leaf mesophyll protoplasts [8]. The remaining candidate RBPs could represent developmental, tissue or cell-specific RBPs, although further studies are required to confirm these results [9]. Interestingly, these studies also identified many proteins with no orthologs in other organisms, no known RBDs, or for which RNA-binding activity was unprecedented. These include the photoreceptors phytochrome A and phototropin1, aquaporins, metabolic enzymes with functions in glycolysis and the citric acid cycle, and enzymes involved in cell redox homeostasis and photosynthesis. Some of these proteins could, therefore, represent 'enigmRBPs' exclusive to plants that possess moonlighting functions in RNA regulation or metabolism [10]. Examples of enigmRBPs such as phosphoglycerate kinase and thioredoxin have been described for other organisms including yeast or human cells [10]. Other proteins identified as putative RNA-binding proteins, which are more widely distributed across eukaryotes, include proteins involved in cytoskeleton and membrane transport, which have been proposed to be involved in RNA transport around the cell and from cell-to-cell.

Perspectives and future directions

The large number of plant RBPs discovered through the application of RNA interactome capture reveals the critical role of RBPs in a cellular context and highlights the importance of RBP-mediated regulation for plant development and physiology. Concordant with the results obtained for other organisms, the subsets of active RBPs are expected to vary depending on cell type, tissue, developmental stage or environmental conditions [10,11]. Therefore, application of RNA interactome capture to different tissues and in different conditions will provide new insights into the dynamic behaviour of plant RBPomes and will help uncover the cues that govern RBP activity, fields that currently remain to be explored. Additionally, with further optimization to capture non-poly(A) RNAs, this technique could be used to identify organellar RBPs as well as the RBPomes of plants and symbionts or pathogens functioning during symbiosis and infection, respectively.

Although all three studies to date have focused on *Arabidopsis,* future studies will expand our knowledge of plant RBPs in crop species, potentially identifying RBPs with critical roles in agricultural traits such as yield, disease resistance or drought tolerance. Because RBPs can regulate the expression of several different genes simultaneously and in response to environmental conditions, RBPs could be of particular interest for plant breeders as targets to increase yield or resistance to stresses. Furthermore, RBPs often follow a modular design with different RBDs and auxiliary domains and some studies have successfully engineered RBPs to modify their RNA-binding activity [12]. Thus, by engineering the specificity, affinity or versatility of a single RBP it may be possible to affect the expression of multiple genes simultaneously with minimal genetic modification.

One of the major impacts of the discovery of the plant RBPome comes to light when put into a wider evolutionary context. Until now, only the RBPomes of *Plasmodium*, *Trypanosoma*, *Leishmania*, yeast, worms, flies, mouse and humans have been described, omitting a crucial group of eukaryotic organisms. Hence, the discovery of plant RBPomes opens the possibility of bridging the evolutionary gap and understanding how RBPomes were shaped during evolution and when and how

individual RBPs evolved. Intriguingly, Reichel *et al*. identified potential novel RBDs in plants such as DUF1296 (Domain of Unknown Function 1296). Experimental validation and structural analysis of these novel RBDs will provide new insights into the molecular basis of RNA-protein interactions, and it will be of interest to determine whether these plant-specific RBDs are linked to plant-specific traits.

The adaptation of RNA interactome capture to the model plant *Arabidopsis thaliana* will also encourage further studies in plants involving UV crosslinking approaches such as crosslinking, immunoprecipitation and sequencing (CLIP-seq) and its variants [13], to identify the RNAs bound by known and newly identified RBPs. This will enable comprehensive determination of RBP-RNA networks occurring in plants and their significance for plant growth, development and responses to environmental conditions.

Concluding remarks

In summary, the studies conducted by Marondedze *et al*., Reichel *et al*. and Zhang *et al*. have provided evidence, for the first time, that RNA interactome capture can be successfully applied to plants. This technique enables system-wide identification of active RBPs and their dynamic regulation in different cell types, tissues, developmental stages and environmental conditions. Additionally, application of RNA interactome capture to plant species represents a promising tool to understand how RBPs and RBPomes were shaped during evolution and a unique opportunity for understanding and engineering RBPs for crop improvement.

Acknowledgements

We thank Anthony A. Millar and Marlene Reichel for providing us with the manuscript Reichel *et al*. (2016) prior to publication. Marcel Bach-Pages is supported by BBSRC (BB/M011224/1) and by the Lorna Casselton Memorial Scholarship at St. Cross College, University of Oxford. Alfredo Castello is funded by the MRC Career Development Award (MR/L019434/1).

References

1 Singh, G. *et al.* (2015) The clothes make the mRNA: past and present trends in mRNP fashion. *Annu. Rev. Biochem.* 84, 583–592

2 Pramanik, S.K. and Bewley, J.D. (1996) Post-transcriptional regulation of protein synthesis during alfalfa embryogenesis: proteins associated with the cytoplasmic polysomal and non-polysomal mRNAs (messenger ribonucleoprotein complex). *J. Exp. Bot.* 45, 1871–1879

3 Castello, A. *et al.* (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406

4 Castello, A. *et al.* (2013) System-wide identification of RNA-binding proteins by interactome capture. *Nat. Protoc.* 8, 491–500

5 Baltz, A.G. *et al.* (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* 46, 674–690

6 Marondedze, C. *et al.* (2016) The RNA-binding protein repertoire of Arabidopsis thaliana. *Sci. Rep.* 6, 29766

7 Reichel, M. *et al.* (2016) In planta determination of the mRNA-binding proteome of Arabidopsis etiolated seedlings. *Plant Cell* 28, 2435–2452

8 Zhang, Z. *et al.* (2016) UV crosslinked mRNA-binding proteins captured from leaf mesophyll protoplasts. *Plant Methods* 12, 42

9 Köster, T. *et al.* (2017) RNA-binding proteins revisited – the emerging arabidopsis mRNA interactome. *Trends Plant Sci.* 22,

10 Beckmann, B.M. *et al.* (2015) The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat. Commun.* 6, 10127

11 Lee, K. and Kang, H. (2016) Emerging Roles of RNA-Binding Proteins in Plant Growth, Development, and Stress Responses. *Mol. Cells* 39, 179–185

12 Wei, H. and Wang, Z. (2015) Engineering RNA-binding proteins with diverse activities. *Wiley Interdiscip. Rev. RNA* 6, 597–613

13 Van Nostrand, E.L. *et al.* (2016) Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* 13, 1–9

Figure 1. Overview of RNA interactome capture.

A. Cells are irradiated with UV light at 254 nm to promote crosslinking between RNAs and proteins that are in intimate contact (1). Next, cells are lysed and mRNAs pulled-down using oligo(dT) magnetic beads (2). After stringent washes, the RNA-protein complexes are recovered and can be analysed using different techniques. Firstly, RNA can be enzymatically digested (3) and the proteins quantitatively analysed by Mass Spectrometry (4) or Western blotting (5). Alternatively, the protein fraction can be enzymatically digested (6) and the RNA analysed by RT-qPCR or RNA sequencing (7).

B. Schematic representation depicting the phylogenetic relationships between the organisms for which RBPomes have been identified by applying RNA interactome capture.

Figure as EPS file